2019

# Acoustic Characteristics Used to Differentiate Speech from Song and Individual Factors that Impact their Effectiveness

Xin Qi
*Western University*

# Acoustic Characteristics used to Differentiate Speech from Song and Individual Factors that Impact their Effectiveness

Xin Qi

Dr. Jessica Grahn

The Department of Physiology and Pharmacology, Schulich School of Medicine & Dentistry, and The Department of Psychology, Brain and Mind Institute, The University of Western Ontario, London, ON, Canada

## Table of Contents

## Abstract

There are many acoustic differences between speech and song, such as frequency range, average fundamental frequency, pitch stability, and rhythmic regularity. Previous studies have shown that musical and linguistic knowledge are recruited differently, but no studies have addressed what specific acoustic features people use to differentiate between speech and song. Our study is designed to determine what acoustic characteristics are used to distinguish speech from song, and to elucidate whether individual factors, such as musical training and tonal language experience, have an effect on these characteristics. In Experiment 1, participants were asked to rank 15 acoustic characteristics according to their importance in differentiating between speech and song. After listening to ambiguous sounding stimuli, participants were asked to re-rank the characteristics. Results showed that melody, beat, and rhythmic regularity were ranked significantly higher ($X^2$=92.69, $p$<0.001) than other characteristics, but these characteristics were not statistically different in their relative rankings to each other. From these results, Experiment 2 had participants categorize sentences as speech or song when we parametrically manipulated the melodic salience of each syllable on a continuum from speech-to-song and from song-to-speech. This was done by manipulating the spoken pitch contour to match the sung pitch contour and vice versa. Decreasing melodic salience resulted in a greater proportion of speech responses, with melodic manipulation and directionality of manipulation having a significant effect on proportion of speech responses ($p$<0.0001). Melodic salience had a greater effect on perception in the song-to-speech direction ($d$=2.67), likely due to a combination of spectral and temporal characteristics affecting stimulus categorization. Musical training and tonal language experience had no effect on response categorization. Results from this study provide insight on the specific cognitive processes used for effective communication in the form of speech and song and contributes to our overall understanding about the way sound is perceived.

**Keywords:** music, language, pitch, rhythm, melody

## Introduction

### *Speech and song in communication*

Speech and song are two universal forms of human communication that share many similarities. Both are similar in their sound production, structural organization, use of acoustic characteristics, and recruitment of sophisticated cognitive and motor processes (Lindblom et al., 2007; Tierney et al., 2013; Patel et al., 2003). Despite this overlap, previous studies have shown that musical and linguistic knowledge are recruited differently, suggesting that listeners without formal musical training and trained musicians alike can easily differentiate between speech and song (Vanden Bosch der Nederlanden et al., 2015). Listeners can vocally imitate pitch changes with greater accuracy when listening to a phrase that is sung, compared to the same phrase that is spoken (Mantell et al., 2013). This demonstrates implicit knowledge regarding pitch, which is recruited more readily in response to a song rather than speech (Mantell et al., 2013). The conclusion that knowledge recruitment is different between the two forms of auditory stimuli does not elucidate the factors listeners use on a daily basis to differentiate between speech and song. It is important to determine these specific factors because the ability to interpret sounds as speech or song may be critical for language development. There is evidence that 2-year-olds apply acoustic knowledge, specifically knowledge about pitch-contours, consistent with their native language, to differentiate novel words (Quam et al., 2009). As such, the extraction of different factors necessary for differentiating between speech and song is a skill present in humans, even at a young age. It is possible that the differences extracted from the message, depending on if it’s perceived as music or language, may impact how that message is encoded and interpreted, Previous studies have not directly identified which factors listeners use.

***Differences in speech and song***

Speech and song are distinct in many ways, including their physical acoustic characteristics, function, the context in which they are used, and emotional impact (Jackendoff, 2009). One way to identify if a phrase is spoken or sung is by evaluating the physical acoustic differences in sound (Vanden Bosch der Nederlanden et al., 2015). Previous research has found that compared to speech, songs tend to have a larger frequency range, higher average fundamental frequencies, greater pitch stability within and between notes or syllables, greater rhythmic regularity, and require greater subglottal pressure to maintain vocalizations. (Lindblom et al., 2007; Patel et al., 2003). These acoustic differences can be classified as temporal or spectral. Temporal characteristics change over time whereas spectral characteristics change within a frequency domain (Bourmans et al., 2007). For example, rhythmic regularity is classified as a temporal acoustic feature because differences in rhythm change over time, whereas pitch stability is an example of a spectral acoustic feature because it changes within a frequency domain rather than across time. Classification of speech and song with the use of acoustic characteristics is of particular interest because although contextual and functional aspects may be useful, these factors alone are unable to explain the breadth of our ability to categorize stimuli as speech or song. There is some evidence that speech and non-speech sounds are processed in different areas of the brain. Many neuroimaging studies demonstrated greater left auditory cortical activity in response to speech (Binder et al., 2000). It is possible that the different responses elicited in the brain may be due to different acoustic features between speech and song (Zatorre et al., 2002). Although there is evidence that acoustical differences exist between speech and song, no studies have addressed what acoustic features people use to differentiate them. Loudness is an example of an acoustic characteristic that differs between speech and song but is not used to distinguish between the two (Lindblom et al., 2007). Considering this, the first aim of our experiment was to determine what acoustic characteristics were used to distinguish speech from song.

***Effects of musical background on perception of speech and song***

Even though it is likely that listeners from all cultures are able to differentiate between speaking and singing, individual differences, such as musical background and language, influence auditory perception (Vanden Bosch der Nederlanden et al., 2015). Musicians have heightened pitch sensitivity and greater knowledge about rhythmic structure (Fujioka et al., 2006; Trainor et al., 2003), but it is unclear how this affects the acoustic characteristics musicians would use to distinguish between music and language, and whether possible differences in perception could be due to low-level acoustic features of the stimulus. There is evidence that the perceptual abilities of non-musicians are as sufficient as the perceptual abilities of musicians, as seen in the speech-to-song illusion—the phenomenon whereby a spoken utterance transforms from speech-to-song when repeated out of context (Vanden Bosch der Nederlanden et al., 2015). However, it remains unclear whether different low-level characteristics are used by musicians compared to non-musicians in order to perceive sound as either speech or song. It is possible that differences in musical ability alter pitch perception because experienced listeners are better able to detect changes in pitch that do not conform to typical pitch patterns in music (Trainor et al., 1992; 1994). Although both musicians and non-musicians exhibit musical knowledge, musicians display greater pitch sensitivity (Vanden Bosch der Nederlanden et al., 2015), which may result in listeners relying more heavily on pitch to distinguish between speech and song.

***Effects of tonal language on perception of speech and song***

Another individual factor to consider is tonal language experience. A language is categorized as tonal if changing the pitch of the word alters its meaning (Wang et al., 2007). In addition to native language experience, stimulus context also contributes to the ability to categorically perceive pitches, but these contextual effects were also more prevalent in native tonal-language speakers (Bidelman et al., 2015). Experience with pitch changes within a linguistic context, such as the expertise obtained from tonal language knowledge, allows a listener to have greater sensitivity to pitch information (Bidelman et al., 2011). Furthermore, the effects of tonal language experience can be seen from the difference in perception of the speech-to-song illusion. Native tonal language speakers perceived this illusion significantly less than non-tonal language speakers because of their effective ability to linguistically categorize various pitch patterns (Jaisin et al., 2016). This finding suggests that language experience alters perception, but it is unclear whether tonal language speakers will use different acoustic features to categorize speech and song. These different perceptual abilities observed in musicians and tonal language speakers provide rationale for investigating the possible effects of individual differences on acoustic features used for stimulus differentiation. Thus, the second aim of our project was to investigate whether individual factors had an effect on the importance of the acoustic characteristics used to distinguish between speech and song.

### ***Effects of acoustic characteristics on perception of speech and song***

In two experiments, we determined 1) which acoustic characteristics people self-reported using to differentiate speech and song, and 2) whether listeners use the reported acoustic features to differentiate speech and song in a categorization experiment. In Experiment 1, we hypothesized that a spectral aspect, such as pitch stability, and a temporal aspect, such as rhythmic regularity, would be important acoustic characteristics used to differentiate between speech and song. Greater pitch stability and rhythmic regularity enhanced the speech-to-song illusion, demonstrating their importance relative to other characteristics (Tierney et al., 2018). It is likely that both features would be rated highly by participants as important for differentiating speech and song, thus, pitch stability and rhythmic regularity were hypothesized to be the most salient. For Experiment 2, we hypothesized that pitch stability affects categorization responses and its effectiveness is impacted by individual differences of the listener. Specifically, we hypothesized that a greater proportion of song categorization will occur in musically trained participants, and less would occur in tonal language speakers when pitch stability is manipulated. This was hypothesized because musicians are sensitive to changes in pitch so it is likely that they require less change in pitch stability to perceive stimuli as sounding song-like. In contrast, tonal language speakers categorize changes in pitch with a linguistic framework (Bidelman et al., 2015), suggesting that changes in pitch stability may be less of an indicator when identifying phrases as sung.

## Methods – Experiment 1

Experiment 1 consisted of an online survey where participants were asked to rank acoustic characteristics from most to least important for distinguishing between speech and song (See Appendix A). A listening quiz was presented, followed by the same acoustic characteristic ranking question.

***Participants***

There were 33 participants who gave consent by indicating that they had read the online letter of information and would like to participate in the study. Results from 3 participants were excluded after failing the attention checks placed throughout the survey (See Procedure). Final results from 30 participants (22 male, 8 female) between the ages of 18-64 (N=5 between 18-24, N=16 between 25-34, N=6 between 35-44, N=2 between 45-54, N=1 between 55-64) were analyzed. All participants were fluent in English. Ten participants indicated that they could speak another language, 8 of whom could speak it fluently. Fluent languages spoken were all non-tonal languages (Tamil, Czech, French, and Malayalam). Out of those 10, 3 participants considered themselves bilingual, defined as speaking a second language 50% of the time or more, and their native language was not English. On average, they began to learn English at the age of 6.33 (range: 6-7 years). Out of the 30 participants, 18 had sung or played an instrument before, and 2 were professional musicians, defined as musicians who are paid to perform or teach music. Participants were recruited through the online platform Amazon Mechanical Turk and received $2.50 as monetary compensation for completion of the survey. The University's Ethics Review Board approved of all materials and procedures.

***Apparatus***

Participants completed the study using their personal computing devices and listened to auditory stimuli in the listening quiz through their own devices as well.

***Stimuli***

Ambiguous stimuli were used for the listening quiz section of this study. Excerpts were chosen based on high ambiguity towards categorizing them as speech or song. Auctioneering, popular rap, and infant directed speech were some examples of stimuli used. Excerpts were 2-5 seconds in length and a total of 13 clips were included in the listening quiz.

***Procedure***

All participants were tested individually through the online platform Amazon Mechanical Turk (mTurk), with the survey hosted on Qualtrics. After participants provided consent, they completed the survey by answering various types of questions (See Appendix A). Part 1 of the survey asked participants to write their answers to questions in short answer format, then code their responses based on a set of pre-determined categories. The rationale behind this was to maintain accuracy in each individual's responses. Following this, participants were asked to rank acoustic characteristics in terms of importance for differentiating between speech and song. Part 2 of the survey consisted of a listening quiz where participants listened to 13 excerpts of ambiguous sounding stimuli and were asked to categorize them as either speech or song. The purpose of this section was to prompt participants to use their musical and language categorization skills they described in the previous section of the survey. Attention checks were placed in this section, which prompted participants to listen to each audio track at least once before proceeding with their answer. After the listening quiz, participants were asked to rank acoustic characteristics again. Lastly, Part 3 of the quiz consisted of demographic questions pertaining to their background, language, music, and dance information. Rankings were analyzed by a Friedman's Test, followed by Wilcoxon Signed Ranks Test using SPSS. Following data analysis, stimuli for Experiment 2 were generated by manipulating one spectral acoustic characteristic, based on the results from Experiment 1.

**Results – Experiment 1**

When participants were asked to categorize their written answer about how music and language differ, 29.7% (n=22) of responses indicated function as the primary category of difference, forming the majority (Table 1). Table 1 also shows that 25.7% (n=19) of respondents chose acoustics as a difference between the two, 22.97% (n=17) chose emotion, 10.81% (n=8) chose context, and 10.81% (n=8) believe that there are no differences between music and language. Participants were asked to select all categories that applied to their written answer, so there was a total of 74 responses for this question from 30 participants. When participants were asked to categorize their answers about the difference between speech and song, 16% of participants chose melody (n=21), 18% chose rhythmic regularity (n=18), and 15% chose pitch height (n=15), forming the top 3 acoustic characteristic choices (Table 1). Total number of responses for this question was 131 from of 30 participants.

Participants were also asked to rank 15 acoustic characteristics according to their importance in differentiating between speech and song. Results showed that melody, beat, and rhythmic regularity were ranked significantly higher ($X^2$=92.69, $p$<0.001), meaning they were considered more important than other characteristics, but these characteristics were not significantly different in their relative rankings to each other (Table 2). Melody, beat, and rhythmic regularity continued to be ranked significantly

**Table 1: Participant responses (n=30) to pre-determined categories when asked to code their written responses to questions regarding the difference between speech and song**

| Question | Response n (%) |
|---|---|
| How would you categorize your answer about how music and language differ? Music and language primarily differ based on: | |
| Function | 22 (29.7) |
| Acoustics | 19 (25.7) |
| Emotion | 17 (22.97) |
| Context | 8 (10.81) |
| There is no difference between music and language | 8 (10.81) |
| Other | 0 (0) |
| How would you categorize your answer about the sound differences between speech and song? Speech and song differ primarily based on: | |
| Melody | 21 (16.0) |
| Rhythmic regularity | 18 (13.7) |
| Pitch height | 15 (11.5) |
| Pitch stability | 14 (10.7) |
| Loudness | 13 (9.9) |
| Variability | 12 (9.2) |
| Pitch range | 10 (7.6) |
| Repetition | 9 (6.9) |
| Feel a beat | 9 (6.9) |
| Duration | 8 (6.1) |
| Other | 2 (1.5) |

*Note:* Participants were asked to select all the categories that applied to their written answer.

**Table 2: Mean rankings (n=30) of acoustic characteristics in terms of importance for differentiating between speech and song, before and after the listening quiz**

| Acoustic Characteristic | Mean Rank Before Quiz | Mean Rank After Quiz |
|---|---|---|
| **Melody** | **4.00** | **3.83** |
| **Rhythmic regularity** | **4.67** | **4.50** |
| **Feel a beat** | **4.43** | **3.40** |
| Pitch height | 6.80 | 6.27 |
| Pitch stability | 5.73 | 5.80 |
| Repetition | 5.73 | 7.50 |
| Pitch range | 5.37 | 5.93 |
| Variability | 5.43 | 6.03 |
| Loudness | 7.33 | 7.23 |
| Duration | 5.67 | 5.27 |
| Other | 10.73 | 10.23 |

*Note:* 1 = most important; 10 = least important. Results showed that melody, beat, and rhythmic regularity (in bold) were ranked significantly higher before ($X^2$=92.69, p<0.001) and after ($X^2$=98.32, p<0.001) the quiz compared other characteristics, but these characteristics were not statistically different in their relative rankings to each other.

higher ($X^2$=98.32, $p$<0.001) than other characteristics after the listening quiz. Although mean rank of the top 3 acoustic characteristics decreased slightly after the quiz, these characteristics were not statically different in relative rankings when compared to each other (Table 2). Before the quiz, melody had the lowest mean rank of 4.00, rhythmic regularity mean rank was 4.67, and beat mean rank was 4.43. After the quiz, melody had a mean rank of 3.83, rhythmic regularity's mean rank was 4.50, and beat mean rank fell to 3.40 (Table 2).

**Methods – Experiment 2**

Results from Experiment 1 determined that the three highest rated spectral and temporal features after the listening quiz were melody, rhythmic regularity, and beat. Melody was ranked as the most salient spectral aspect used to differentiate speech from song. As such, we manipulated how salient the melody of utterances was by making the pitch contour of each spoken syllable match it's sung counterpart and vice-versa. This resulted in changing the pitch stability of syllables within a sentence. We manipulated melodic salience to investigate whether changes in melody are what listeners use to distinguish between speech and song.

***Participants***

Thirty-one undergraduate students (10 male, 21 female) between the age of 19-26 currently attending Western University were recruited by word-of-mouth communication and through Western University's undergraduate psychology participant pool. Participants received course credit for participation. Those who were not eligible for course credit were compensated $5 per half-hour of testing. All participants were fluent in English. Twenty-three participants indicated that they could speak another language, 18 of whom could speak it fluently. Fluent languages spoken included tonal languages (N=4) and non-tonal languages (N=14). Tonal languages spoken were Mandarin (N=1), Cantonese (N=1), and Taiwanese (N=1). Eighteen participants have taken private music lessons for over 5 years (5-18 years). All participants provided informed consent prior to the experiment. After completing the experiment, participants filled out a demographic questionnaire which included questions regarding their hearing, music, and language abilities. All information obtained from the questionnaire remained confidential. The University's Ethics Review Board approved all materials and procedures prior to testing.

***Apparatus***

All participants were tested individually in a quiet room using a Windows 7 Dell Precision laptop running Intel Core i7. Stimuli were presented to participants through Sennheiser HDR 160 headphones at a comfortable listening volume as determined by the participant (mean volume = 35%).

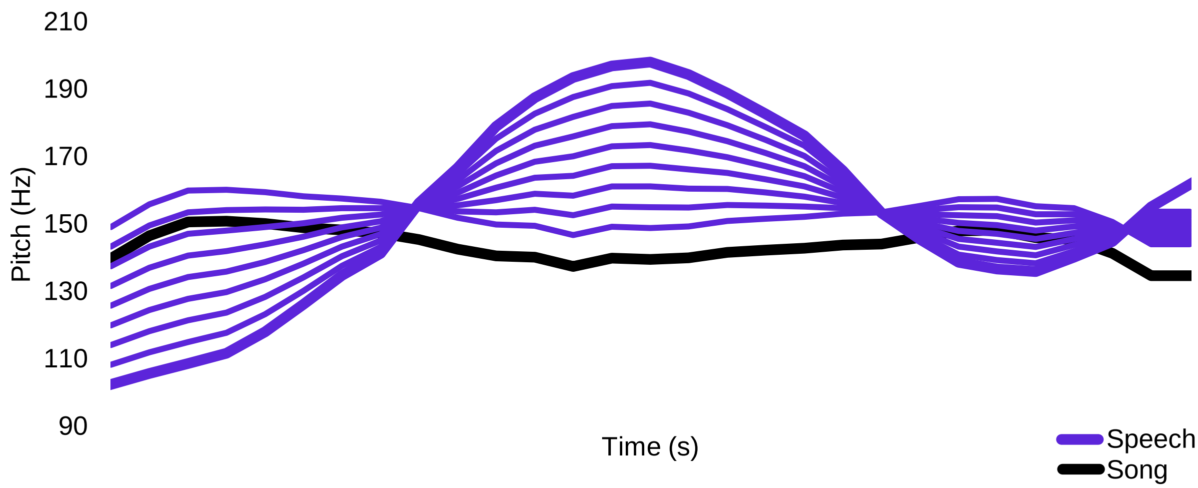
***Stimuli***

A corpus of 24 excerpts of spoken sentences and 24 matched excerpts of the same sentences, but sung, were obtained from a previous study in preparation and used with permission (Vanden Bosch der Nederlanden, 2016). Using Praat, a computer program for the analysis and manipulation of sound, pitch and duration tiers for the corpus of stimuli were identified. Pitch information for each syllable in the sound file were extracted, such as start time, end time, duration, average pitch, standard deviation, maximum, and minimum of fundamental frequency. This information was used to determine 13 matched sentences as good candidates for pitch manipulation. Good candidates were identified as sentences where the difference between average standard deviation of pitch for spoken and sung sentences were greater than 1. Melodic salience of a spoken sentence was altered to produce 8 intermediate sentences, with each one progressing towards the final sung sentence in a step-wise manner (Figure 1A). This was done by isolating single syllables in a sentence and determining equal step-sizes by finding the difference between the syllable's particular pitch tier and the average pitch of that syllable. Step sizes were applied such that each manipulation from speech-to-song constructed a new list of pitch values that was manipulated around the original spoken syllable's average pitch, and became progressively more like the sung contour, to generate a continuum that became more song-like. This procedure was repeated for each syllable in the 13 good candidates identified from the original corpus. Manipulation constructs were used in a Praat script to generate a total of 130 sentences where manipulations resulted in a continuum from speech-to-song. Melodic salience of a sung sentence was also altered to produce 8 intermediate sentences, with each one progressing towards the final spoken sentence in a step-wise manner using the same procedure (Figure 1B). A total of 130 sentences were generated in a continuum from song-to-speech. Spoken and sung original contours were also run through the re-synthesis procedure in Praat so that all stimuli were reproduced through the same method. As such, duplicates of spoken and sung contours were removed for each sentence due to excessive distortion, resulting in a corpus of stimuli with 234 sentences, each 2 seconds in length.

***Procedure***

The experiment was run on E-prime. Participants were given verbal instructions by the experimenter to listen to a sentence and categorize it as either speech or song

A




B

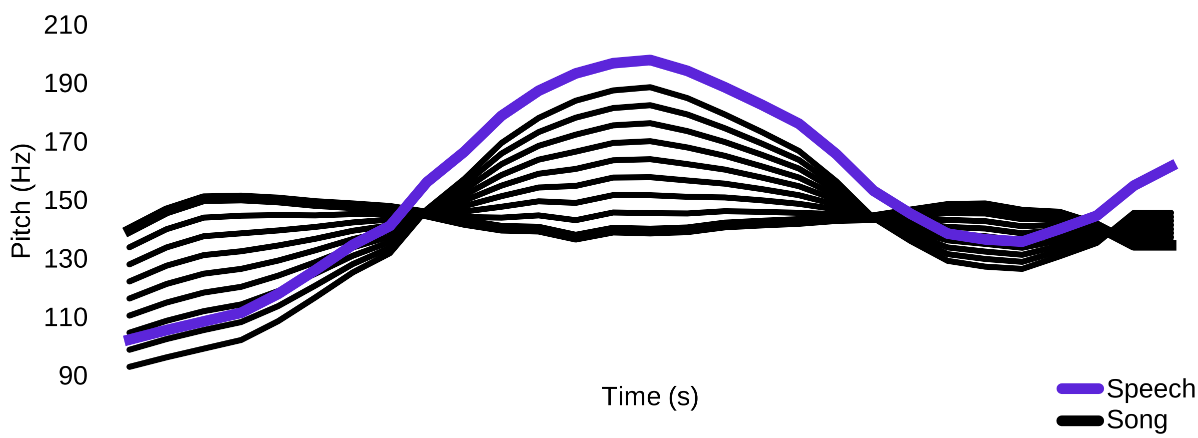



**Figure 1.** Tracings of the various manipulations of pitch contour (Hz) of a single syllable "rice" over time (s). The spoken (A) or sung (B) pitch contour of the syllable "rice" was manipulated to produce 8 intermediate step-wise pitch contours, with each one progressing towards a song-like contour (A) or a speech-like contour (B). (A) shows direction manipulation from speech-to-song. (B) shows direction manipulation from song-to-speech.

immediately after listening to it. Participants were instructed to press the “speech” button (s key) if it sounded like a spoken sentence and press the “song” button (l key) if it sounded like a sung sentence. Keyboard buttons were labelled with their respective designation. Participants were not able to re-listen to the sentence. After this, participants were asked to provide a subjective rating of how confident they were about their categorization, from 1 (Not confident at all) to 5 (Very confident) by pressing the appropriate number key on the keyboard. On-screen instructions were written prior to the start of the experiment. This procedure was repeated for all 234 sentences, presented to participants in a random order, and were counterbalanced. Lastly, participants were asked to fill out a questionnaire of demographic questions which pertained to their education, language experience, and musical background. Participants (N=8) were considered tonal language speakers if they indicated they could speak a tonal language on a fluency rating between 1 (Very fluent) to 6 (Slightly fluent), and participants (N=16) were considered musicians if they had a self-reported musical skill rating of greater than 4 on a scale from 1-6 and had 5 or more years of private music lessons. A mixed logistic regression analysis was conducted in which slope and intercept coefficients were compared to determine a model between melodic salience and direction of manipulation that best described the data, followed by a t-test and Cohen’s d test to examine the significance and effect size of each manipulation level on speech or song categorization. The same statistical analysis was done to determine the effect of melodic salience and direction on confidence rating.

## Results Experiment 2

### *Stimulus Categorization Results*

We entered participants stimulus categorization responses into a logistic regression with manipulation, direction, manipulation by direction interaction, musicianship, musical skill, and tonal language as predictors. In 6 models, we examined the contribution of each factor to the overall model. Model 3 led to the largest amount of variance explained ($R^2$=0.342), while still resulting in a significant increase in fit from the previous model ($p$<0.0001), indicating that model 3 is the best model for our results and is the greatest predictor of speech or song choice (see Table 3). Model 3 shows the effect of the variables manipulation, direction, and the manipulation by direction interaction on proportion of speech ratings (Figure 2). Our model with musical skill, model 5, also reached significance, but the size of the beta coefficient for a musical skill rating of 3 suggests that this rating alone is what is driving the significant result, thus, results from this model are not reported. Models 4 and 6, which show the effect of musical background and tonal language experience respectively, did not result in a significant increase in fit ($p$=0.2419; $p$=0.1909), suggesting that these factors do not predict speech or song choice (Figures 3 and 4).

Given the significant interaction term in model 3, we performed a paired-samples t-test to determine whether manipulation had different effects depending on the direction of the manipulations (speech-to-song or song-to-speech). Both directions of manipulations had significant effects (*p's*<0.001), but a Cohen's d test to compare their effect size determined that the direction of manipulations going from song-to-speech ($d$=2.67) is greater than from speech-to-song ($d$=1.49). Although melodic salience

**Table 3: Significant logistic regression model for speech or song categorization**

| Model 3 | Variable | AIC | Chi-square | DF | p |
|---|---|---|---|---|---|
| | Manipulation | 9623.6 | 679.48 | 9 | <0.0001 |
| | Direction | 8630.8 | 703.35 | 1 | <0.0001 |
| | Manipulation:Direction | 8182.3 | 425.91 | 9 | <0.0001 |

*Note:* Akaike Information Criterion (AIC) is an estimator of the quality of the model relative to other models. AIC provides an estimate for the amount of data lost by the proposed model, so the smaller the number, the higher accuracy of the model.

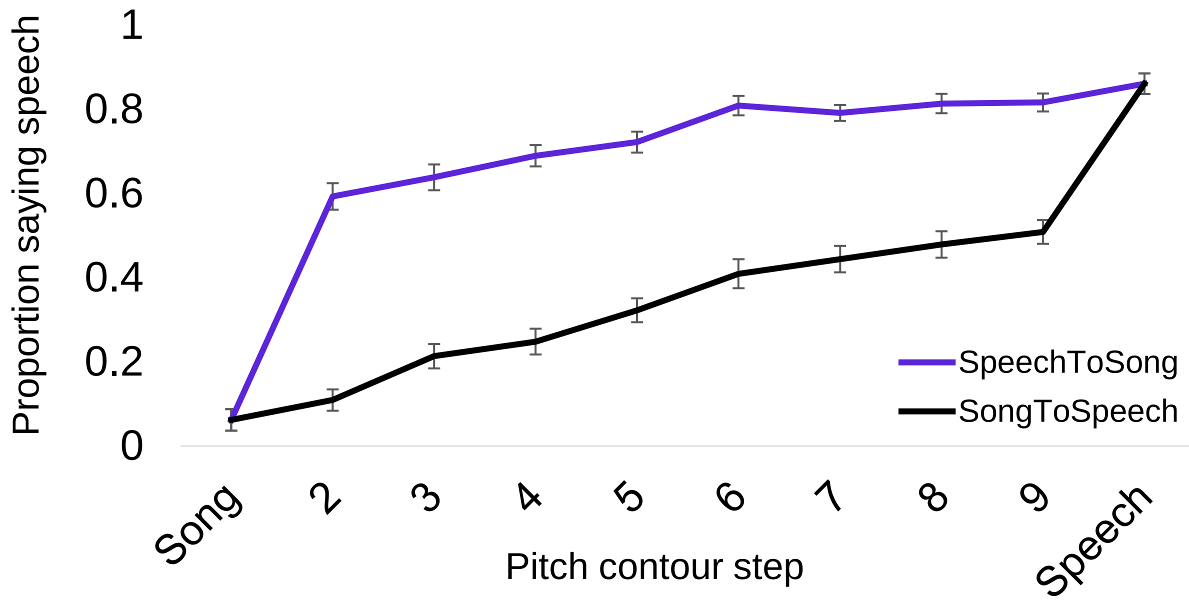



**Figure 2.** Participants categorized auditory stimuli, presented in a random order and counterbalanced, as either speech or song immediately after listening. Data shown are mean (N=31) proportion of speech responses ± SEM across all 10 melodic salience manipulation steps after listening to sentences manipulated in both the speech-to-song direction, and the song-to-speech direction. Logistic regression analysis showed a significantly accurate model based on manipulation by direction interaction on stimuli categorization ($R^2$=0.342). Specifically, melodic salience manipulation had a greater effect size when manipulated in the song-to-speech direction (*d*=2.67).

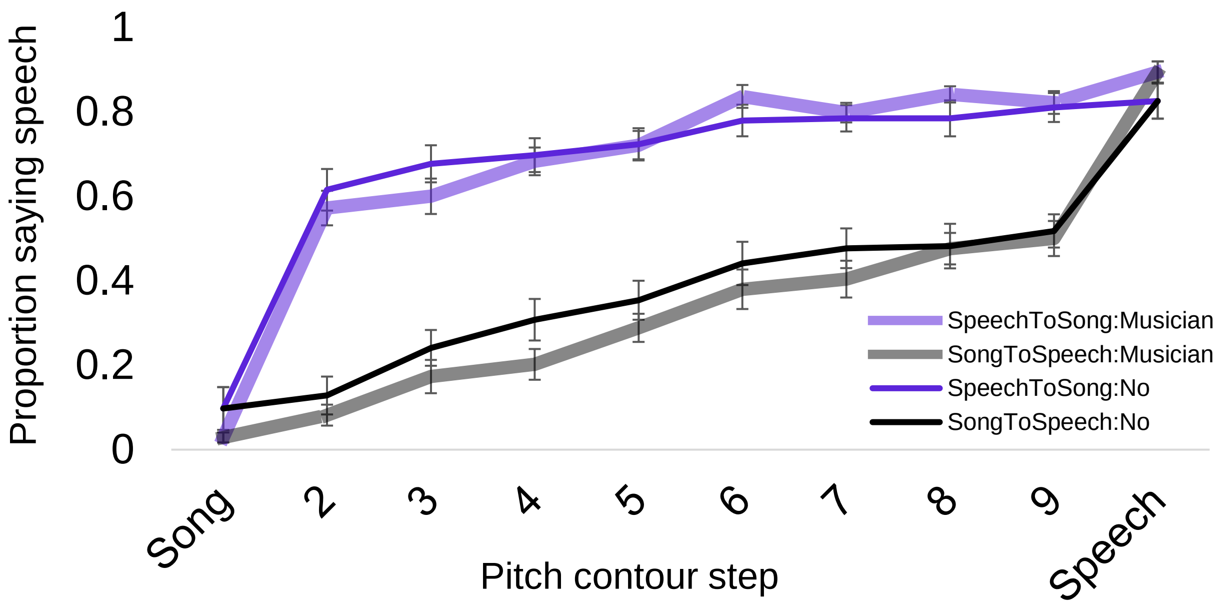



**Figure 3.** Musicians and non-musicians categorized auditory stimuli, presented in a random order and counterbalanced, as either speech or song immediately after listening. Data shown are mean (N=31) proportion of speech responses ± SEM across all 10 melodic salience manipulation steps after listening to sentences manipulated in both the speech-to-song direction, and the song-to-speech direction. Logistic regression analysis showed a non-significant increase in fit of the model based on the effect of musicianship on stimuli categorization ($R^2$=0.401).

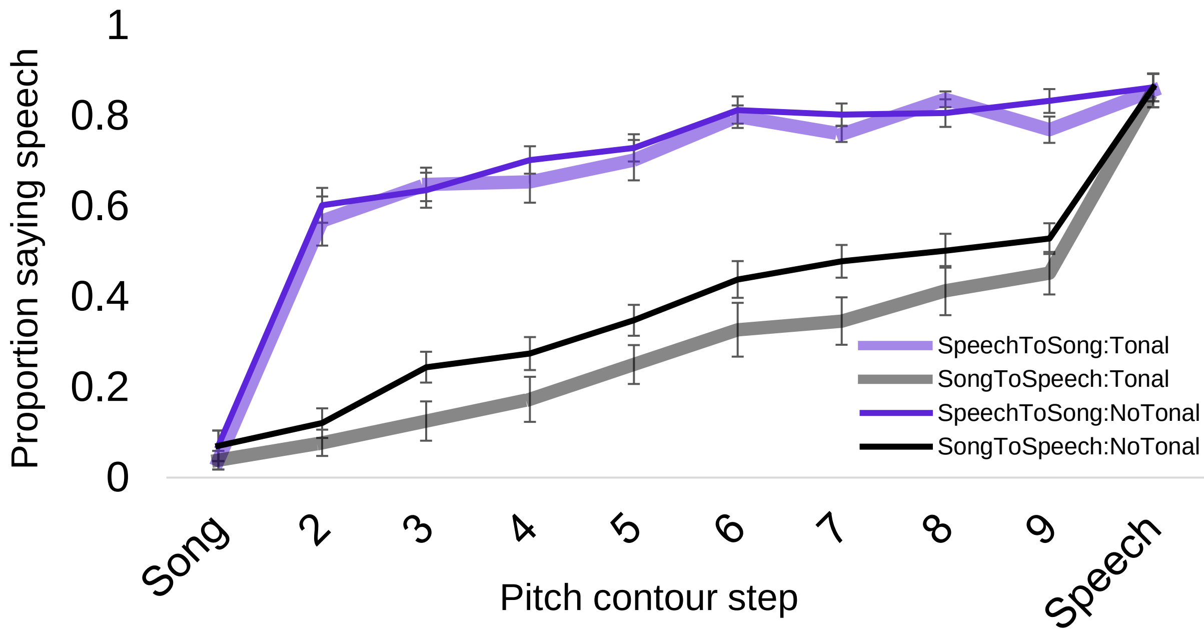



**Figure 4.** Tonal language and non-tonal language speakers categorized auditory stimuli, presented in a random order and counterbalanced, as either speech or song immediately after listening. Data shown are mean (N=31) proportion of speech responses ± SEM across all 10 melodic salience manipulation steps after listening to sentences manipulated in both the speech-to-song direction, and the song-to-speech direction. Logistic regression analysis showed a non-significant increase in fit of the model based on the effect of tonal language experience on stimuli categorization ($R^2$=0.401).

manipulation had significant effects in both directions, it had a greater effect size when pitch stability was manipulated from a sung to a spoken utterance.

### *Confident Rating Results*

We entered participants confidence ratings into a logistic regression with manipulation, direction, manipulation by direction interaction, musicianship, musical skill, and tonal language as predictors. Confidence ratings were made binary with not confident rated between 1-3 and confident rated between 4-5. In 6 models, we examined the contribution of the addition of each factor to the overall model. Model 2 led to the largest amount of variance explained ($R^2$=0.0573), while still resulting in a significant increase in fit from the previous model, indicating that model 2 is the best model for our results and is the greatest predictor of confidence rating (see Table 4). Model 2 shows the effect of the variables manipulation and direction on confidence ratings. Models 3-6 did not increase the fit of the model. Since model 3 was not significant, there was no interaction between manipulation and direction, meaning the effect of manipulation for both directions was not significantly different (Figure 5).

## Discussion

### *Experiment 1*

One spectral aspect (melody) and 2 temporal aspects (beat and rhythmic regularity) were ranked significantly higher than other acoustic characteristics but were not differentially ranked between themselves. These results were consistent both before and after the listening quiz, indicating the perceived importance of melody, beat, and rhythm relative to other acoustic characteristics, but none considered more important than the others. Although both beat and rhythm are temporal features of sound, the beat is the

**Table 4: Significant logistic regression model for confidence ratings**

| Model 2 | Variable | AIC | Chi-square | DF | P |
|---|---|---|---|---|---|
| | Manipulation | 8470.9 | 320.536 | 9 | <0.0001 |
| | Direction | 8446.3 | 26.519 | 1 | <0.0001 |

*Note:* Akaike Information Criterion (AIC) is an estimator of the quality of the model relative to other models. AIC provides an estimate for the amount of data lost by the proposed model, so the smaller the number, the higher accuracy of the model.

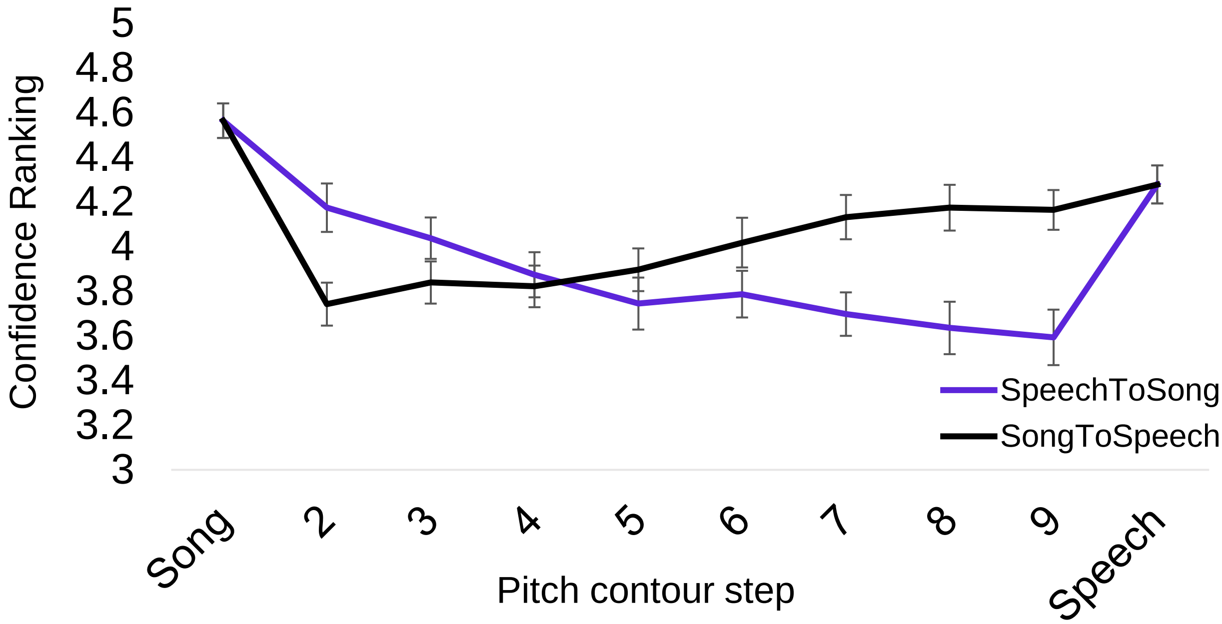



**Figure 5.** Auditory stimuli were presented in a random order, and participants categorized each sentence as either speech or song, and then were prompted to rank their confidence on their categorization, with 1 = Not Confident and 5 = Very Confident. Data shown are mean (N=31) confidence ranking ± SEM across all 10 melodic salience manipulation steps after listening to sentences manipulated in both the speech-to-song direction, and the song-to-speech direction. Logistic regression analysis showed a significantly accurate model based on direction ($R^2$=0.0573). There was no interaction between manipulation and direction.

pulse felt throughout the piece, whereas rhythm refers to the pattern of the notes themselves. Contrary to what was hypothesized, melody was the spectral characteristic that participants determined as the most salient, rather than pitch stability. A possible explanation is due to the heightened memorability of short, repetitive phrases of a melody, which are more prevalent in songs rather than speech (Janssen et al., 2017). Although, it is also possible that participants may have ranked melody significantly higher than other characteristics because songs tend to have greater variability in pitch range while maintaining greater syllable-level pitch stability (Lindblom et al., 2007). Variability in pitch range can contribute to a prominent melody, a factor that participants may have ranked highly, contrary to what we predicted, since pitch stability may be too technical of an aspect for listeners to grasp. In addition, beat and rhythmic regularity were also ranked significantly higher than other acoustic characteristics possibly because participants recognized the importance of considering temporal features when evaluating the difference between speech and song, consistent with findings from Tierney and colleagues (2018), which showed that rhythmic regularity was vital for enhancing the speech-to-song illusion.

These results contribute to our understanding of how people think about, and perhaps distinguish between speech and song. However, the use of surveys and questionnaires only yields self-reported results which does have limitations. Self-reported data provides an understanding that on average, people think that melody, beat, and rhythm play an important role in perception, but it is not known whether manipulating these features would result in a change in percept from speech-to-song or vice versa, which provided rationale to manipulate melodic salience in Experiment 2.

### *Experiment 2*

We manipulated melodic salience in 8 steps from speech-to-song and song-to-speech and found that melodic salience significantly affected the proportion of speech responses, with a greater effect size in the song-to-speech direction. We anticipated a difference in categorization responses after manipulating melodic salience because speech and song have different pitch stabilities and pitch contours (Lindblom et al., 2007). Although results from Experiment 1 prompted us to manipulate melodic salience rather than pitch stability, which is what was predicted in our hypothesis, our results are still consistent with what was hypothesized. Listeners do use spectral aspects, specifically melodic salience manipulated by mimicking the sung or spoken contours, to differentiate speech from song, confirming that acoustic differences between speech and song are useful for differentiating these two modes of human communication. Surprisingly, the direction of manipulation also had a significant impact on categorization as either speech or song. Melodic salience had a greater effect on perception in the song-to-speech direction, meaning that each manipulation step that began with a song contour was more likely to be categorized as sounding speech-like the more speech-like the contour became. Comparatively, in the speech-to-song direction, as the speech contour was manipulated to sound more song-like—that is, although participants' ratings changed with the degree of the manipulation, responses were still mostly consistent with the stimulus sounding like speech. Since the direction of manipulation had a significant effect, this may provide evidence that spectral features may not be the only characteristics affecting perception. Temporal features, such as rhythmic regularity, which was identified as an important characteristic in Experiment 1, may explain why direction significantly impacted perception. If melodic salience were the only aspect used to differentiate between speech and song, direction of manipulation should have negligible effects, since all other characteristics of the original sentence remained unmanipulated. However, since direction did have an effect, a characteristic from the original spoken or sung sentence also influenced perception. It is likely that a combination of temporal and spectral features is used to discriminate between speech and song (Falk et al., 2014). When speech pitch contours were manipulated to become more song-like, the irregular rhythm pattern associated with spoken sentences remained consistent for all manipulations. Although pitch tracings became more song-like, combined with irregular rhythms, the listener's perception to whether the sentence is speech or song was split between two conflicting signals.

This explanation is consistent with confidence ratings. Although model 3 was not significant, indicating that there was no interaction between manipulation and direction for confidence ratings, direction itself did significantly affect confidence levels. Overall, participants were significantly less confident in their responses in the speech-to-song direction, and significantly more confident in the song-to-speech direction. In the speech-to-song direction, confidence ratings trended towards a decrease with each manipulation step, likely because of the ambiguity caused by conflicting spectral and temporal signals. In contrast, when manipulating melodic salience in the song-to-speech direction, sung sentences tended to be more rhythmically regular than speech, but regular intervals can also occur in speech (Rathcke et al., 2015). When pitch contours were manipulated to sound more speech-like, participant rankings trended towards increasing confidence levels in their categorization with each manipulation step. Thus, the direction of manipulation had an effect on overall confidence levels.

Another possible explanation for the significant effect that melodic salience and direction manipulation had on stimulus categorization is that when manipulating melodic salience in the speech-to-song direction, even though the pitch contour became more song-like, the pitch itself still followed the spoken contour, and thus, was not fitted to the typical Western musical scale that one would expect when listening to songs. Similar to the rhythmic aspect, this conflicting signal may also have contributed to responses that were more consistent with the stimulus sounding like speech. This explanation is supported by a previous study which found that pitch discrimination was significantly better when music-specific pitch representations violated Western musical scale structure, showing that participants are using their knowledge of musical scale structure to listen to songs (Vanden Bosch der Nederlanden et al., 2015).

Our findings also reveal that individual differences, such as musical background and tonal language experience, do not impact the effectiveness of melodic salience on stimulus categorization. Although our model for musical skill's effect on categorization response was significant, the beta coefficient for the third rating was an outlier compared to the rest of the ratings, suggesting that this is what resulted in significance. As such, it is likely that this result, although significant, would not be replicable in repeated studies, since there is no logical explanation as to why rating musical skill as a 3 would contribute to greater proportion of speech ratings. In addition, musicianship and tonal language experience did not affect categorization response, since model 4 and 6 were not significant. This finding was contrary to our hypothesis, which stated that musicians and tonal language speakers would have different perception. It is possible that despite musicians having greater knowledge about pitch and rhythm structures, their perceptual abilities in terms of stimulus categorization are comparable to non-musicians, similar to conclusions from Vanden Bosch der Nederlanden and colleagues (2015). Similarly, although tonal language speakers are familiar with pitch changes in a linguistic context, enhanced knowledge about acoustic characteristics do not affect perception (Bidelman et al., 2011). An alternate explanation could be due to power. Out of 31 participants, only 8 spoke tonal languages, and all 8 tonal language speakers were also musically trained, so it is possible that there was a confounding effect. Future studies should use a balanced sample size for tonal language speakers and eliminate the possible confounding effects of musical ability.

A possible limitation with our study design was distortion of stimuli due to excessive transformations of the pitch contour. Future studies should replicate our protocol and add a pilot study where participants rate sound quality to see if it had an effect on categorization responses. In addition, future studies should focus on investigating the effect of temporal characteristics, such as rhythmic regularity or beat, on categorization responses and determine if we use a combination of spectral and temporal features to distinguish between speech and song. These studies would elucidate whether an interaction between spectral and temporal acoustic characteristics is more prevalent in a particular direction of manipulation.

Our research aligns with a larger body of work which aims to understand how humans are able to communicate effectively with each other. We determined that melodic salience is a factor that listeners use to differentiate between speech and song. This recognition of subtle differences in melodic salience between different forms of acoustic input may be an important feature necessary for language development, so an interesting direction to take our research would be to investigate how specific processes in the brain function differently depending on if sound is perceived as speech or song. Once we gain a better understanding about how our perceptual abilities function, we can develop interventions to potentially treat language developmental and communication disorders.

**Acknowledgements**

I would like to thank everyone in the Grahn Lab for this wonderful experience and making this study possible. Specifically, I would like to thank Dr. Jessica Grahn for giving me the amazing opportunity to work in her lab and for introducing me to this fascinating field of research. In addition, I would also like to give an extra special thank you to Christina Vanden Bosch der Nederlanden for her continuous guidance and support throughout my study. I am so grateful for all of her help, particularly with the creation of the stimuli and statistical analysis, and for her helpful words of wisdom and motivation. Lastly, I would like to express my gratitude for the Department of Physiology and Pharmacology, specifically Dr. Angela Beye and Dr. Anita Woods for organizing the PhysPharm 4980 thesis course and enriching the educational experience of all students enrolled. Funding was provided by NSERC.

## References


1. **Bidelman GM, Gandour JT, Krishnan A.** Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *J Cogn Neurosci* 23: 425-434, 2011.
2. **Bidelman GM, Lee CC.** Effects of language experience and stimulus context on the neural organization and categorical perception of speech. *NeuroImage* 120: 191-200, 2015.
3. **Binder J, Frost J, Hammeke T, Bellgowan P, Springer J, Kaufman J, Possing J.** Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex* 10: 512-528, 2000.
4. **Bourmans T, Theunissen FE, Poirier C, Van Der Linden A.** Neural representation of spectral and temporal features of song in the auditory forebrain of zebra finches as revealed by functional MRI. *Eur J Neurosci* 26(9): 2613-2626, 2007.
5. **Deutsch D, Henthorn T, Lapidis R.** The illusory transformation of speech to song. *Journal of the Acoustical Society of America*. 129(4): 2245-2252, 2011.
6. **Falk S, Rathcke TV, Dalla BS.** When speech sounds like music. *J Exp Psychol Hum Percept Perform* 40(4): 1491-1506, 2014.
7. **Fujioka T, Ross B, Kakigi R, Pantey C, Trainor LJ.** One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain* 129: 2593-2608, 2006.
8. **Jackendoff R.** Parallels and nonparallels between language and music. *Music Perception* 26(3): 195-204, 2009.
9. **Jaisin K, Suphanchaimat R, Figueroa Candia MA, Warren JD.** The Speech-to-Song Illusion is Reduced in Speakers of Tonal (vs. Non-Tonal) Languages. *Front Psychol* 7: 662, 2016.
10. **Janssen B, Burgoyne JA, Honing H.** Predicting variation of folk songs: a corpus analysis study on the memorability of melodies. *Front Psychol* 25(8): 621, 2017.
11. **Lindblom B, Sundberg J.** The human voice in speech and singing. In: *Springer Handbook of Acoustics* 669-712, 2007.
12. **Mantell JT, Pfordresher PQ.** Vocal imitation of song and speech. *Cognition* 127: 177-202, 2013.
13. **Patel AD, Daniele JR.** An empirical comparison of rhythm in language and music. *Cognition* 87(1): B35-45, 2003.
14. **Quam C, Swingley D.** Phonological knowledge guides two-year-olds' and adults' interpretation of salient pitch contours in word learning. *J Mem Lang* 62(2): 135-150, 2009.
15. **Rathcke, TV, Smith RH.** Speech timing and linguistic rhythm: on the acoustic bases of rhythm typologies. *J Acoust Soc Am* 137(5): 2834, 2015.
16. **Tierney A, Dick F, Deutsch D, Sereno M.** Speech versus song: multiple pitch-sensitive areas revealed by a naturally occurring musical illusion. *Cereb Cortex* 23(2): 249-254, 2013.
17. **Tierney A, Patel AD, Breen M.** Acoustic foundations of the speech-to-song illusion. *J Exp Psychol Gen* 147(6): 888-904, 2018.

18. **Trainor LJ, Shahin A, Roberts LE.** Effects of musical training on the auditory cortex in children. *NY Acad Sci* 999: 506-513, 2003.
19. **Trainor LJ, Trehub SE.** A comparison of infants' and adults' sensitivity to Western Musical Structure. *J Exp Psychol Hum Percept Perform* 18(2): 394-402, 1992.
20. **Trainor LJ, Trehub SE.** Key membership and implied harmony in Western tonal music: Developmental perspectives. *Perception & Psychophysics* 56(2): 125-132, 1994.
21. **Vanden Bosch der Nederlanden CM, Hannon EE, Snyder JS.** Everyday musical experience is sufficient to perceive the speech-to-song illusion. *J Exp Psychol Gen* 144(2): e43-e49, 2015.
22. **Vanden Bosch der Nederlanden CM, Hannon EE, Snyder JS.** Finding the music of speech: Musical knowledge influences pitch processing in speech. *Cognition* 143: 135-140, 2015.
23. **Vanden Bosch der Nederlanden C, Hannon EE, Snyder JS.** Development of music-specific pitch processing. *J Acoust Soc Am* 140: 3390, 2016.
24. **Wang L, Ambikairajah E, Choi EHC.** A novel method for automatic tonal and non-tonal language classification. *IEEE International Conference* 352-355, 2007.
25. **Zatorre RJ, Belin P, Penhune VB.** Structure and function of auditory cortex: music and speech. *Trends Cogn Sci* 6(1): 27-46, 2002.

**Appendix A: Experiment 1 Survey Questions**

# Speech vs. Song

Welcome to the study. You will be responding to questions about everyday sounds. Sometimes you will listen to sounds and rate them based on what they sound like. Your responses will be collected as part of a study being conducted at Western University in London, Ontario, Canada. Before we get started, please read the letter of information for a detailed overview of the study by clicking this text.

Please indicate your response to continue:

- ○ I have read the letter of information and would like to participate in the study
- ○ I would not like to participate in the study

**Part 1**

What is the difference between music and language?

______________________________________________________________

______________________________________________________________

______________________________________________________________

______________________________________________________________

______________________________________________________________

Think of someone speaking and someone singing. What are the sound features, or the physical properties of sound, that differ between speech and song?

______________________________________________________________

______________________________________________________________

______________________________________________________________

______________________________________________________________

______________________________________________________________

How would you categorize your answer about how music and language differ? Please select all that apply to your written answer

- ▢ Acoustics: I wrote about the way they sound (pitch and rhythm) e.g., "language is less dynamic than music"
- ▢ Function: I wrote about the different ways I engage with music/language e.g., "language is important for communicating ideas"
- ▢ Emotion: I wrote about the way they affect your mood e.g., "music helps calm me down"
- ▢ Context: I wrote about the context in which they are heard/used e.g., "music is often heard at parties or live events"
- ▢ There is no difference between music and language
- ▢ Other (please specify) ________________________________________________

How would you categorize your answer about the sound differences between speech and song? Please select all that apply to your written answer

- ▢ Pitch height: higher pitch vs. lower in pitch
- ▢ Pitch stability: one has a more consistent pitch whereas the other doesn't hit a specific pitch (glides around)
- ▢ Rhythmic regularity: regular vs. irregular patterns unfolding in time
- ▢ Melody: sequences of pitches are more prominent in one than the other
- ▢ Repetition: words, phrases, or notes are repeated vs. little or no repetition
- ▢ Pitch range: the range of pitches is smaller in one compared to the other
- ▢ Variability: many different notes vs. notes are mostly the same
- ▢ Loudness: overall one is louder or softer than the other

☐ Feel a beat: you could clap/tap along with one, but not the other

☐ Duration: words, phrases, or notes are longer in one than the other; more silences in one compared to the other

☐ Other (please specify)

________________________________________________

Please rank these characteristics in terms of how important they are for differentiating speech and song.

______ Pitch height: higher vs. lower in pitch
______ Pitch stability: one has a more consistent pitch whereas the other doesn't hit a specific pitch (glides around)
______ Rhythmic regularity: regular vs. irregular patterns unfolding in time
______ Melody: sequences of pitches are more prominent in one than the other
______ Repetition: words, phrases, or notes are repeated vs. little or no repetition
______ Pitch range: the range of pitches is smaller in one compared to the other
______ Variability: many different notes vs. notes are mostly the same
______ Loudness: overall one is louder or softer than the other
______ Feel a beat: you could clap/tap along with one, but not the other)
______ Duration: words, phrases, or notes are longer in one than the other; more silences in one compared to the other
______ Other (please specify)

Have you given much thought about the differences between speech and song, prior to this survey?

○ None at all

○ A little

○ A moderate amount

○ A lot

○ A great deal

Do you think it is easy to tell when someone is singing vs. when someone is speaking?

○ Extremely easy

○ Moderately easy

- Slightly easy
- Neither easy nor difficult
- Slightly difficult
- Moderately difficult
- Extremely difficult

If you have any comments, please feel free to write them here:

________________________________________________________________

________________________________________________________________

**Part 2**

Listening Quiz. For this section, listen to the audio clips presented and categorize it as "speech" or "song". After this, rank the confidence of your categorization on a scale from 1= Not confident at all, 5 = Very confident

Listen to this track

Would you categorize the audio track as speech or song?

1. Speech
2. Song

How confident are you about your categorization?

- 1 (Not confident at all)
- 2 (Not too confident)
- 3 (Neither confident or not-confident)
- 4 (Confident)
- 5 (Very confident)

**Questions above were repeated for 12 other audio tracts.

Now that you've listened to these music and language clips, how would you rank the following acoustic characteristics in terms of how important they are for differentiating speech and song?

______ Pitch height: higher pitch vs. lower in pitch
______ Pitch stability: one has a more consistent pitch whereas the other doesn't hit a specific pitch (glides around)
______ Rhythmic regularity: regular vs. irregular patterns unfolding in time
______ Melody: sequences of pitches are more prominent in one than the other
______ Repetition: words, phrases, or notes are repeated vs. little or no repetition
______ Pitch range: the range of pitches is smaller in one compared to the other
______ Variability: many different notes vs. notes are mostly the same
______ Loudness: overall one is louder or softer than the other
______ Feel a beat: you could clap/tap along with one, but not the other
______ Duration: words, phrases, or notes are longer in one than the other; more silences in one compared to the other
______ Other (please specify)

If you have any comments, please feel free to write them here:

____________________________________________________________

____________________________________________________________

____________________________________________________________

____________________________________________________________

____________________________________________________________

**Part 3**

Now we are going to get a little background information about you so we can better understand your responses.

Age:

- ◯ Under 18
- ◯ 18 - 24
- ◯ 25 - 34
- ◯ 35 - 44
- ◯ 45 - 54
- ◯ 55 - 64
- ◯ 65 - 74
- ◯ 75 - 84
- ◯ 85 or older

Gender:

______________________________________________________________

Year in school:

- ◯ Less than high school
- ◯ High school graduate
- ◯ Some college
- ◯ 2 year degree
- ◯ 4 year degree
- ◯ Professional degree
- ◯ Doctorate

What is your race? Please check all that apply.

- ☐ White
- ☐ Puerto Rican
- ☐ Mexican, Mexican-American, or Chicano
- ☐ Cuban
- ☐ Asian Indian
- ☐ Korean
- ☐ Native Hawaiian
- ☐ Samoan
- ☐ Black/African American
- ☐ Chinese
- ☐ Vietnamese
- ☐ Guamanian/Chamorro
- ☐ American Indian/Alaska Native
- ☐ Filipino
- ☐ Japanese
- ☐ Other Asian: ________________________________________
- ☐ Other Pacific Islander: ________________________________________

☐ Other Spanish/Hispanic/Latino: ________________________________________________

☐ Other race: ________________________________________________

Mother's highest education level?

- ○ No H.S. diploma
- ○ H.S. diploma
- ○ Some college
- ○ 4-year college degree
- ○ Graduate school degree
- ○ Technical school

Father's highest education level?

- ○ No H.S. diploma
- ○ H.S. diploma
- ○ Some college
- ○ 4-year college degree
- ○ Graduate school degree
- ○ Technical school

Did you learn English from birth?

- ○ Yes
- ○ No

What is your native language?

________________________________________________________________

How old were you when you began learning English?

________________________________________________________________

Do you speak any other languages? (Languages other than English)

○ Yes

○ No

We will ask you about your top 3 other languages (languages other than English). List one of the other languages that you speak.
________________________________________________________________

What is your competency?

○ N/A

○ Beginner

○ Intermediate

○ Advanced/Fluent

Do you speak a second non-english language?

○ Yes

○ No

List the second other language that you speak.

________________________________________________________________

What is your competency?

- ◯ N/A
- ◯ Beginner
- ◯ Intermediate
- ◯ Advanced/Fluent

Do you speak a third non-english language?

- ◯ Yes
- ◯ No

List the third other language that you speak.

________________________________________________________________

What is your competency?

- ◯ N/A
- ◯ Beginner
- ◯ Intermediate
- ◯ Advanced/Fluent

Do you consider yourself bilingual?

- ◯ Yes
- ◯ No

What do you consider your dominant/main language?

________________________________________________________________

What percentage of the time do you speak your dominant/main language? i.e. 50%? 30%?

________________________________________________________________

Have you lived in any country outside of Canada for more than 6 months?

- ◯ Yes
- ◯ No

Where?

________________________________________________________________

How long?

________________________________________________________________

Describe your exposure to music and/or dance there:

________________________________________________________________

Where were you born?

________________________________________________________________

Where was your mother born?

________________________________________________________________

Where was your father born?

________________________________________________________________

Did you ever sing or play an instrument?

- ◯ Yes
- ◯ No

How would you describe yourself as a musician?

◯ Occasional musician (less than weekly practice/participation)

◯ Recreational musician (weekly practice or recreational playing/performance)

◯ Serious amateur musician (extensive commitment to practice and/or recreational music activity)

◯ Professional musician (paid to perform and/or teach music)

Type of music practiced (Classical/Jazz/Folk/etc.)?

______________________________________________________________

What instrument(s) have you played? Voice can be included as an instrument.

______________________________________________________________

Have you ever played an instrument and/or sung in an ensemble? (i.e. school band, orchestra, choir etc.)?

◯ Yes

◯ No

Type of ensemble (check all that apply):

- ☐ School band
- ☐ Private Institute Band
- ☐ Self-Arranged Band/Orchestra Ensemble
- ☐ School Orchestra
- ☐ Private Institute Orchestra
- ☐ School Choir
- ☐ School Theatre Group
- ☐ Self-Arranged Choir Ensemble
- ☐ Other: ________________________________________________

Beginning at what age?

________________________________________________________________

No. of years?

________________________________________________________________

Have you ever taken private music lessons?

- ○ Yes
- ○ No

Beginning at what age?

________________________________________________________________

No. of years?

________________________________________________________________

Solo or group lessons? (please describe if group):

____________________________________________________________

Are you currently taking private lessons?

- ◯ Yes
- ◯ No

Instrument:

____________________________________________________________

How many days per week are the lessons?

____________________________________________________________

How many hours per day are the lessons?

____________________________________________________________

How often do you play/sing music on a weekly basis?

- ◯ 1 day
- ◯ 2-3 days
- ◯ 4-5 days
- ◯ 6-7 days

On average, how many hours per day do you play music? (Practice and recreationally)

____________________________________________________________

Have you performed or taught music professionally? (i.e. for pay)

- ◯ Yes
- ◯ No

How many years?

__________________________________________________________________

Do you dance (recreationally, formally, etc.)

- ◯ Yes
- ◯ No

How would you describe yourself as a dancer?

- ◯ Occasional Dancer (less than weekly dancing for fun or practice)
- ◯ Recreational Dancer (weekly practice or recreational dance)
- ◯ Serious Amateur Dancer (extensive commitment to practice and recreational dance activity)
- ◯ Professional Dancer (paid to perform and/or teach dance)

Type(s) of dance practiced? Please check all that apply.

- [ ] Folk
- [ ] Ballet
- [ ] Hip-Hop
- [ ] Middle Eastern
- [ ] Contra-dance
- [ ] Jazz
- [ ] Asian
- [ ] Ballroom
- [ ] Flamenco/Latin
- [ ] Contemporary
- [ ] Tap
- [ ] Lyrical
- [ ] Other: ____________________________________________

What age did you start dancing?

__________________________________________________________

No. of years?

__________________________________________________________

Have you ever participated in formal dance lessons?

- ○ Yes
- ○ No

Beginning at what age?

______________________________________________________________

No. of years?

______________________________________________________________

Are you currently taking dancing lessons?

- ○ Yes
- ○ No

What type of dance?

______________________________________________________________

How many days per week are the lessons?

______________________________________________________________

How many hours per day are the lessons?

______________________________________________________________

How often do you dance on a weekly basis?

- ○ 1 day
- ○ 2-3 days
- ○ 4-5 days
- ○ 6-7 days

On average, how many hours per day do you dance? (Practice and recreationally)

______________________________________________________________

Have you danced professionally? (i.e. for pay)

- ◯ Yes
- ◯ No

How many years?

______________________________________________________________

Can you read music?

- ◯ Yes
- ◯ No

Have you ever taken music courses at the university level?

- ◯ Yes
- ◯ No

Which course(s)?
______________________________________________________________

Do you have formal training in music theory (classes or self-taught)?

- ○ Yes
- ○ No

If yes, how many years?

- ◯ 0.5
- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4-6
- ◯ 7+

Do you have absolute pitch? (i.e. if someone played a note on the piano, could you name the note without looking)

- ◯ Yes
- ◯ No
- ◯ Don't know

On average, how many hours per week do you listen to music?

________________________________________________________________

What types of music do you listen to?

________________________________________________________________

When you listen to music, do you primarily listen to the lyrics, or to the melody?

- ◯ Lyrics
- ◯ Melody
- ◯ Both

When you read in your head, do you hear a voice in your head "speaking" the words that you are reading?

- ○ Yes
- ○ No
- ○ I'm not sure

Have you gotten goosebumps/shivers from listening to music before?

- ○ Yes
- ○ No

Are any of your family members musicians?

- ○ Yes
- ○ No

If yes, which family members (what is their relationship to you i.e. mother, brother, cousin, etc.)?

________________________________________________________________

Are any of your family members dancers?

- ○ Yes
- ○ No

If yes, which family members (what is their relationship to you i.e. mother, brother, cousin, etc.)?

________________________________________________________________

During what other activities do you like to listen to music? Please list

________________________________________________________________